



A novel application of deep learning with image cropping: a smart city use case for flood monitoring

MISHRA, B.K., THAKKER, D., MAZUMDAR, Suvodeep <<http://orcid.org/0000-0002-0748-7638>>, NEAGU, D., GHEORGHE, M. and SIMPSON, S.

Available from Sheffield Hallam University Research Archive (SHURA) at:

<http://shura.shu.ac.uk/26010/>

This document is the author deposited version. You are advised to consult the publisher's version if you wish to cite from it.

Published version

MISHRA, B.K., THAKKER, D., MAZUMDAR, Suvodeep, NEAGU, D., GHEORGHE, M. and SIMPSON, S. (2020). A novel application of deep learning with image cropping: a smart city use case for flood monitoring. *Journal of Reliable Intelligent Environments*, 6 (1), 51-61.

Copyright and re-use policy

See <http://shura.shu.ac.uk/information.html>



A novel application of deep learning with image cropping: a smart city use case for flood monitoring

Bhupesh Kumar Mishra¹ · Dhavalkumar Thakker¹ · Suvodeep Mazumdar² · Daniel Neagu¹ · Marian Gheorghe¹ · Sydney Simpson³

Received: 3 June 2019 / Accepted: 21 January 2020 / Published online: 24 February 2020
© The Author(s) 2020

Abstract

Event monitoring is an essential application of Smart City platforms. Real-time monitoring of gully and drainage blockage is an important part of flood monitoring applications. Building viable IoT sensors for detecting blockage is a complex task due to the limitations of deploying such sensors in situ. Image classification with deep learning is a potential alternative solution. However, there are no image datasets of gullies and drainages. We were faced with such challenges as part of developing a flood monitoring application in a European Union-funded project. To address these issues, we propose a novel image classification approach based on deep learning with an IoT-enabled camera to monitor gullies and drainages. This approach utilises deep learning to develop an effective image classification model to classify blockage images into different class labels based on the severity. In order to handle the complexity of video-based images, and subsequent poor classification accuracy of the model, we have carried out experiments with the removal of image edges by applying image cropping. The process of cropping in our proposed experimentation is aimed to concentrate only on the regions of interest within images, hence leaving out some proportion of image edges. An image dataset from crowd-sourced publicly accessible images has been curated to train and test the proposed model. For validation, model accuracies were compared considering model with and without image cropping. The cropping-based image classification showed improvement in the classification accuracy. This paper outlines the lessons from our experimentation that have a wider impact on many similar use cases involving IoT-based cameras as part of smart city event monitoring platforms.

Keywords Image classification · Deep learning · DCNN · IoT sensors · Drainage blockage

1 Introduction

Image classification methods group images into predefined class labels after analysing image features. Deep convolutional neural networks (DCNN), with several layers and a large number of nodes in each layer, have been extensively used for image classification [1–5]. In a deep-learning-based model, the performance of the model is dependent upon the number of training datasets used to train the classification model [6]. To achieve higher classification accuracy, image classification with deep learning requires to be trained with a large number of images [7, 8]. However, in recent years, the use of DCNN has been growing in many application areas where a significant number of training datasets are not always available [2]. Deep learning-based image classifications have shown higher accuracy in several application areas such as extracting characteristic features to distinguish objects and classify images into different class labels. Also, applying

✉ Dhavalkumar Thakker
D.Thakker@bradford.ac.uk

Bhupesh Kumar Mishra
b.mishra@bradford.ac.uk

Suvodeep Mazumdar
s.mazumdar@sheffield.ac.uk

Daniel Neagu
d.neagu@bradford.ac.uk

Marian Gheorghe
m.gheorghe@bradford.ac.uk

Sydney Simpson
sydney.simpson@bradford.gov.uk

¹ University of Bradford, Bradford, UK

² University of Sheffield, Sheffield, UK

³ City of Bradford Metropolitan District Council, Bradford, UK

DCNN, higher accuracy has been achieved for recognising objects in an image [9].

With access to affordable GPU resources, the use of deep learning has shown to outperform some of the conventional methods of image classification in many cases [1]. Despite the promising results on applying deep learning techniques for image classification, significant challenges exist in its application domains with limited access to training datasets. Without sufficient training data, the accuracy of the learning phase of the classification model can suffer from the problem of underfitting, ultimately leading to a degraded classification model [5]. Apart from the limited availability of datasets for training, the wider context and the number and types of distinct objects within images play a crucial role in image classification [2, 7, 8]. Often, the same image can be classified into different classes based on the context and objects. Besides this, the context for different classes can be quite different which adds additional complexity in classifying into multi-labels [10].

We encountered the issue of limited image datasets and identification of context in the use case of flood monitoring as part of an EU Interreg project SCORE¹ (Smart Cities + Open Data Re-use). The project involves nine cities, and three universities to explore the use of IoT and big data technologies for solving challenges facing cities in water, environment, and mobility themes. As a part of the SCORE project, our research focuses on addressing challenges relevant to water theme, in particular, flood monitoring. The level of blockage of gullies and drainage systems is one of the important parameters for local authorities and city councils to monitor in real time. IoT sensors for detecting blocked gullies and drainage are notoriously hard to build because of the limitations of deploying them in situ. Hence, as a replacement, we implemented a DCNN classifier for image classification with a low number of image dataset collected from Google images, YouTube videos and other sources to classify blockage incidents. Whilst our longer term ambition is to develop in situ IoT sensors that can automatically detect drain blockages as soon as they occur, this paper focuses on the scalable approach to classifying images by exploring the regions of interest from images achieved by applying image cropping at data pre-processing stages. We applied edge removal using image cropping to localise the region of interest in images before images are utilised for the training and validation phases. The experimental results showed that image classification using image cropping at data pre-processing stage has higher accuracy in comparison to the classification accuracy of image classification without cropping. We also compared the performance of the proposed approach with support vector machine (SVM) classifiers. The comparative analysis

showed that the proposed classification approach has higher accuracy than the SVM classifiers.

The major contribution of this paper is in applying deep learning-based image classifier to monitor drain and gully blockages using image capturing devices. Use of deep learning-based image classification in smart cities, with flood monitoring as a use case, shows that this proposed approach has high potential as an alternative approach to the complex sensor-based real-time monitoring systems. This is because the installation of physical sensor devices in some geographical areas is not feasible. Use of camera imagery and deep learning-based image classifier for the detection of drain and gully blockage using a reusable dataset is a unique contribution of this work. The second contribution of this work comes from the use of image cropping at data pre-processing stage to reduce any form of noise present at the edges of the image. The image cropping techniques are advantageous as it provides a more focused region of interest within an image to improve feature extraction from images. Since our source of images is from a camera focused and pointing towards a gully or drainage, this approach has resulted in increased accuracy in our experimentation. The concept of image cropping at data pre-processing stage as in our work offers a useful comparison for other applications where the characteristic features from images are also focused in the central region. The third contribution of this work is the image dataset that we have systematically curated from scratch since none exists in the public domain so far.

The rest of the paper is organised as follows: Sect. 2 reviews image classification approaches. Section 3 describes the concept of deep learning, image cropping and image classification, whereas Sect. 4 describes the experimental design, data generation and result analysis. We conclude the paper and describe future work in Sect. 5.

2 Literature review

Several methods have been applied for image classification tasks such as k-nearest neighbours, maximum likelihood minimum distance, decision tree, and logistic regression [10]. Most of these methods suffer from the “curse of dimensionality” [11] as there occur a number of dimensions within image dataset, making dimensionality reduction an essential task in image analysis. In general, the reduction of the dimension of image dataset is performed at the initial stages to improve the classification performance. Reduction in dimensionality to fit the input image data into small-scale feature extraction has been explored using techniques such as principal component analysis (PCA) [11, 12], pairwise constraint discriminant analysis, and non-negative sparse divergence (PCDA-NSD) [13]. Non-negative matrix factorization (NMF) [14], transformation [15] and band selection [16] have also been applied

¹ <https://northsearegion.eu/score/>.

for feature extraction. In addition to these, the spectral-spatial feature-based classification (SSFC) framework [17] has been applied as well for feature extraction. The SSFC framework used both dimensionality reduction and deep learning techniques for spectral and spatial feature extraction.

Image classification has been performed using algorithms such as decision tree (DT) [18], random forest (RF) [19, 20], K-nearest neighbour (KNN) [21, 22] and support vector machine (SVM) [10, 23, 24]. DT algorithms used rules to interpret the appropriate class of the input datasets [18] whereas RF algorithms were used for non-normally distributed and high-dimensional data and were also more stable than DT algorithms [25]. K-nearest neighbour (KNN) classifiers had also been used for image classification. A KNN classifier uses characteristic features and similarity between the features to classify images into k class labels. The SVM method has been reported to have promising results in pattern recognition for supervised classification [23]. SVMs were initially defined as binary classifiers, and their use for multi-label classifications had been more challenging. SVM has been used for image classification as a supervised classifier [10]. In SVM, an optimal hyperplane is generated to separate the classes using the training dataset that lies at the edges and centroid. For smaller datasets, SVM is one of the highly suitable classifiers as support vectors can identify an optimal separating hyperplane [26].

Over many years, two common approaches (pixel-based and object-based analysis) have been implemented for image classification problems [27]. It was found that object-oriented classification had higher accuracy in comparison to pixel-based classification [28] when a supervised maximum likelihood classification algorithm was applied for object-oriented classification and region-growing multi-resolution cropping with nearest neighbour classifier was applied for pixel-based image classification. However, in general, image classification accuracy between pixel-based and object-based classification approaches was statistically not significant when the same machine learning algorithms were used [29]. In other words, there were no reported advantages of favouring one approach over another when the same algorithm had been implemented.

Other alternative image classification approaches had also been used. Multi-object detection approaches had been applied to classify images which constitute multiple objects representing multiple scenarios [30, 31]. To classify multi-object images [30], one classifier had been trained for each object to be analysed and a score based on probability was provided to test images. These classified outputs were combined with the probabilities of the individual classes to classify images. Traditionally, pixel-based classification methods have been used for image analysis in many applications while the integration of the concept of object-based analysis into the image classification process showed an

increment in the accuracy [27, 32]. This implies that to reach higher precision in image classification, the integration of additional information within an image needs to be analysed. In another approach to detect multiple objects, a number of trained detectors were applied to determine characteristic visual features from images [31]. Also, semantic representation is also utilised for multi-object image classification. In such an approach, objects rather than features had been used as attributes to characterise images [33].

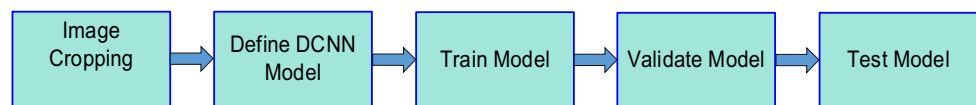
The accuracy of image classification tasks depends on the volume of data available for training and validation. Some models [7, 8] have used thousands of images to train the classification model. While typical supervised models need large volumes of training image dataset to learn their model parameters, specific problem-based image classifiers have been trained with only a limited number of sample images. Deep learning models have had reasonably higher accuracy for image classification, where large image dataset was not available for training [4, 34]. Simple image classification tasks have also been observed with relatively higher accuracy using smaller datasets.

In recent years, intelligent system approaches such as neural network models have also been used to improve performance on image classification tasks [1, 2, 35, 36]. Deep belief network (DBN) has been used for spectral-spatial classification of hyperspectral images by applying a hybrid framework with PCA hierarchical learning-based feature extraction, and logistic regression. The DBN model has used a single-layer restricted Boltzmann machine and multilayer deep network-based models to learn the shallow and deep features of hyperspectral data, correspondingly [36]. In the DBN model, linear regression was applied to classify images based on the extracted features. Residual learning framework was used for training of networks [37] to increase the training efficiency by explicitly fitting stacked layers to the desired underlying mapping instead of direct mapping. DBN with restricted Boltzmann machine (RBM) was applied [38] for image classification by extracting features of images.

Over the years, deep convolutional neural network (DCNN) with multiple layers of neurons has been commonly used for image classification. Each neuron in DCNN extracts a different level of non-linear characteristic features from an image. DCNN was applied to classify image data using large volumes (1.2 million high-resolution images) of training data in imageNet LSVRC-2010 contest into 1000 different classes [8]. Not only with large image dataset but DCNN has also been applied with smaller numbers (few thousands) of training image data for classification [1]. DCNN has been found to be a useful tool for image classification tasks.

A wide range of image classification approaches has been applied in diverse applications where images were classified using either pixel-based analysis or object-based analysis. It has been found that image classification models were mostly

Fig. 1 Building blocks of image classifier



being implemented on static image classification. However, in real-time scenarios such as monitoring flooding where the camera acts as a sensor, the classification needs to be applied on real-time streamed images. A number of alternative methods have been implemented for image classification, but in recent years deep learning approaches have been widely used for image classification. It has also been observed that the classification of an image using deep learning gives higher accuracy, which is highly desirable for multi-label, real-time image classification tasks.

3 Deep learning and image cropping: our approach

We propose a deep learning-based image classifier where images are categorised into classified images. These unclassified images can be classified into either “fully blocked” or “partially blocked” or “no blockage” classes. The building blocks of image classifier are shown in Fig. 1. To start with, image cropping is applied at the data pre-processing stage where images are cropped with 10% edge removal from all sides. At the second stage, DCNN parameters such as type of activation functions, number of classes, batch size, error rate, number of iterations are defined. Once the models’ parameters are defined, the model is trained and validated with a training image dataset. After training and validation, the model has been tested with test images. For the improvement in classification accuracy, the model has been trained and validated iteratively by adjusting the model’s parameters such as error rate, batch size, and the number of iterations.

3.1 Image cropping

Image cropping gives the freedom to pick the selected region of interest within the image for the analysis. Image cropping allows to explore target objects or concentrate on a single target region [27]. The process of image cropping into a targeted region is one way to explore the contextual features in terms of area of interest of an image. We analysed that most of the images in our image dataset have non-essential objects such as vehicles, human, and houses. With a low number of image dataset for training, finding a distinct context within an image is challenging, and may lead to poor performance in terms of classification accuracy.

It is commonly observed that while recording an event or capturing images of an incident, such as drainage blockage through a camera, it is typical to have the centre of the

image in focus [39]. This implies the objects and other elements towards the edges, as we move away from the centre, are less significant to analyse the context of an image. Thus, centrally focused image analysis to build an image classifier can be an effective method for image classification. In our image classification approach, we have applied image cropping technique to locate the central area of an image that potentially reduces the impact of objects at the edges of the image.

Traditionally, visual attention to analysing an image can be either region based or object based [40]. The region-based approach directs the attention to the areas of interest, whereas the object-based approach directs the attention to an object or set of objects. In our model, image cropping confines the regions of interest leaving out non-essential objects of images at the edges. Figure 2a is an example of a drain blockage image in its originally captured form whereas Fig. 2b is the same image after applying image cropping. In the resulting image, with 10% pixels cropping applied, it is observed that there has been no alteration in the focused region. Analysing these two images before and after image cropping, it is visibly clear that prominence of some of the non-essential objects or context such as houses and parked van in the image is minimised. We applied pixel-wise cropping to crop image edges. Doing so, one of the challenges is to find the appropriate cropping proportion. We have implemented different proportions ranging from 5 to 20% cropping of edge’s pixels from all four sides to find the best-suited cropping proportions. The deep learning-based image classifier model has been trained with a dataset of cropped images for all these cases. The accuracies and losses are compared for each case to find the most feasible cropping proportion.

3.2 Region of interest

In general, humans indirectly apply structural knowledge to analyse an image during manual classification processes by considering contextual information along with the information about the shape of objects and spatial relations between image regions [32]. Objects, environment and region of interest within an image are critical for feature extraction and hence the classification task. Region of interest is the section within the image which possesses the characteristic features of the image that define the class of image. The region of interest within an image varies with the application areas, image contents, image orientation and classification goals. Considering the supervised monitoring, it has been observed that the edges of images more often do not contain more

Fig. 2 **a** Image without cropping. **b** Image after cropping



characteristic features. Therefore, minimising the impact of edges will enhance the extraction of characteristic features from an image. The removal of edges from image concentrates the feature extraction within the region of interest. In this work, we focused on the feature extraction from an image by creating a centrally focused region of interest. The region of interest within an image has been obtained by removing pixels from the edges of the captured or monitored image during the data pre-processing stage.

4 Experimental design

4.1 Dataset

The quality of training data, a sufficient number of representative images, and diversity in image types representing different classes are the key elements of developing an effective image classifier. Accessing such a dataset to build an image classifier for flood monitoring has been a challenging task since there has been no dedicated and publically available image source for the domain under investigation. To overcome this, we have chosen public platforms such as Google Images and YouTube to collect representative images involving drains and gullies. The motivation for using images from these data sources is as these images have been uploaded or made available by the public which could be anticipated in realistic situations.

Multiple keywords such as “drainage blockage”, “gully blockage”, “drainage blockage and flooding”, “sewage blockage”, “drainage overflow” were applied to search images and videos. It was observed that the majority of search results came up with only a small number of related images which is one of the limitations of gathering image from public platforms. In addition, the public platform images are noisy, distorted and blurred. In general, public platform data are not completely related to a single area. So, after every search, we applied manual cleaning to keep only the relevant images, i.e. images which had gully/drainage and blockage representing three classes we have in our experimentation.

Table 1 Image dataset breakdown

Class	Training	Validation	Test
Fully blocked	230	100	60
Partially blocked	260	100	60
No blockage	230	100	60

The same process also removed noisy, distorted, blurred and out-of-context images. This process was applied during all the searches to generate the required dataset. Overall, nearly 3000 images were manually inspected with the aforementioned process. After further inspection and removing irrelevant images, 1200 images were prepared for the DCNN model.

The next step was labelling of sample images for the training dataset by five experts in drainage division from the City of Bradford Metropolitan District Council. These experts are involved in the day-to-day monitoring of gullies as part of the council and highways. The selection of classification categories was according to the need of the application, and the utility of these classes to represent the severity of the problem as experts see it as a standard in their work. Randomly selected images from the dataset were shown to the experts and were asked to provide binary and multi-label on each image. In a binary labelling, “blocked” and “no blockage” class labels were considered, whereas in multi-class labelling, images were classified into three class labels: “fully blocked”, “partially blocked” and “no blockage”. We applied the majority poll to decide the class label of the image when there were conflicts in experts’ opinion. Furthermore, with analysis of other image classification model designed with higher accuracy, the dataset is grouped into training data, validation data and test data in proportion 60%, 25% and 15%, respectively, as shown in Table 1.

4.2 Experimentation and results

A number of algorithms have been applied for image classification task in different domains. We have applied a deep learning algorithm to build an image classifier for our flood

Table 2 Simulation setup parameters

Class mode	Categorical
Loss	Categorical_crossentropy
Optimizer	Optimizer_rmsprop(lr = 2e-5)
Steps per epoch	100
Epochs	100
Batch size	10
Validation steps	100
CNN network	VGG Net

monitoring application. In addition to applying deep learning algorithm for model training, validation and testing, we have additionally applied image cropping to enhance classification accuracy. In this section, we present the parameters and design of our proposed image classification model. Table 2 shows the simulations set up for the training of our image classification model. These parameters are set after applying multiple ranges or experimenting various parameters such as batch sizes ranging from 5 to 15, steps per epochs and validation steps ranging from 50 to 100 during training and validation phases along with different optimiser values. The accuracies on classifying test images are analysed on each alternative setup. During the analysis, we observed that setting the higher batch size had lower accuracy on classification in comparison with the lower batch size. Therefore, a lower batch size has been set for this model. Among the number of alternative setups, the simulation parameters with the highest accuracy on classification are selected for our classification model.

4.2.1 VGG net

VGG net consists of convolutional layers with a uniform architecture. This network is defined using 3×3 convolutional layers stacked on top of each other in increasing depth. VGG Net is used for the feature extraction of images. The network has convolution kernel, max-pooling, flatten and dense layers. The detailed design of the VGG with the number of layers, number of nodes and activation functions of our deep learning model is summarised in Table 3.

4.3 Result analysis

The image classifier has been trained, validated and tested for the image dataset with and without carrying out image cropping. The process of image cropping is carried out by varying image edge cropping proportions. While applying image cropping, the major question has been “what proportion of image edges can be cropped?” There is no accepted standard in terms of deciding the right proportion of image cropping without losing image’s characteristic features. To

Table 3 Model summary

Conv2D	input	(None, 150,150,3)
	output	(None, 148,148,64)
Activation	input	(None, 148,148,64)
	output	(None, 148,148,64)
MaxPooling2D	input	(None, 148,148,64)
	output	(None, 74,74,64)
Conv2D	input	(None, 74,74,64)
	output	(None, 72,72,32)
Activation	input	(None, 72,72,32)
	output	(None, 72,72,32)
MaxPooling2D	input	(None, 72,72,32)
	output	(None, 36,36,32)
Conv2D	input	(None, 36,36,32)
	output	(None, 34,34,64)
Activation	input	(None, 34,34,64)
	output	(None, 34,34,64)
MaxPooling2D	input	(None, 34,34,64)
	output	(None, 17,17,64)
Flatten	input	(None, 17,17,64)
	output	(None, 18496)
Dense	input	(None, 18496)
	output	(None,64)
Dense	input	(None, 64)
	output	(None,3)
Activation	input	(None,3)
	output	(None,3)

decide the right proportion, we implemented a different range of image cropping (5%, 10%, 15% and 20% image pixels) to decide the best proportion value. For the evaluation of the best cropping proportion, accuracy and loss values for training and validation are recorded for each case, as shown in plots from Figs. 3, 4, 5 and 6. The accuracy and loss val-

Fig. 3 Loss and accuracy plots on model training and validation with 5% cropping

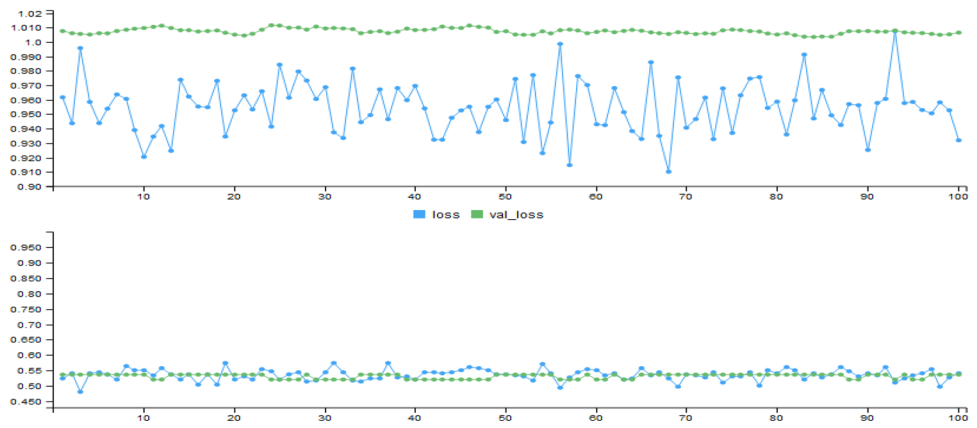


Fig. 4 Loss and accuracy plots on model training and validation with 10% cropping

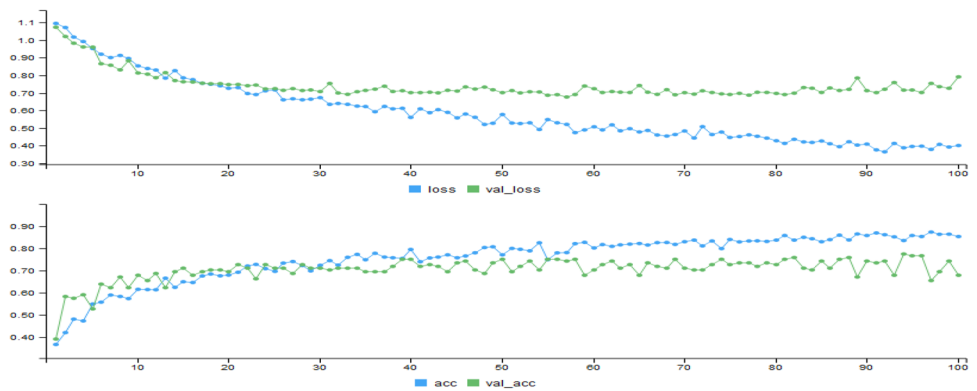


Fig. 5 Loss and accuracy plots on model training and validation with 15% cropping

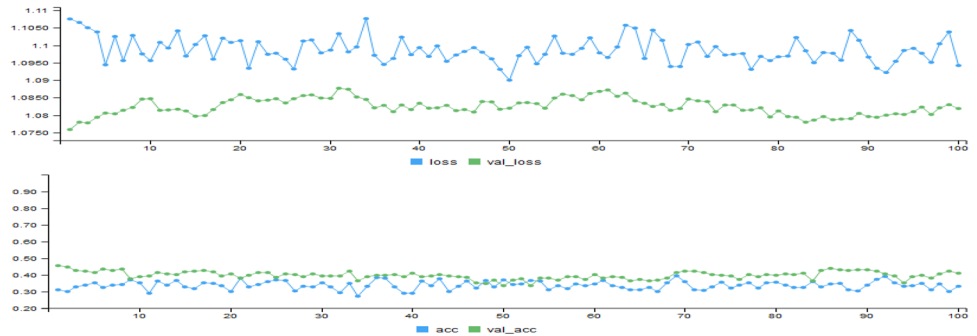
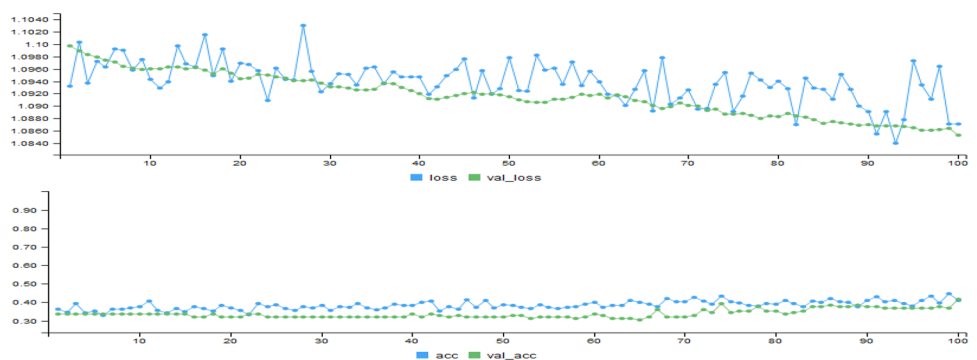


Fig. 6 Loss and accuracy plots on model training and validation with 20% cropping



ues of the different cropping range (as listed in Table 4) are compared to select the best cropping condition.

Accuracies and loss plots during the training of the model with different image cropping range are analysed to choose the best possible image cropping proportion. With 5% of

Table 4 Loss and accuracy values at different cropping values after 100 epochs

Parameter	Training accuracy	Validation accuracy	Training loss	Validation loss
5% cropping	0.54	0.54	0.930	1.010
10% cropping	0.76	0.65	0.4	0.8
15% cropping	0.35	0.45	1.085	1.09
20% cropping	0.4	0.4	1.088	1.086

Table 5 Performance comparison between binary and categorical class mode classification

Model/parameter	Accuracy (training)	Accuracy (test)	Loss
Deep leaning with cropping-binary	0.74	0.68	1.37
Deep leaning with cropping-categorical	0.76	0.66	0.4

edge cropping, it is observed that accuracy values for both the training and validation of the model are almost uniform throughout the 100 epochs. There has not been any noticeable improvement in accuracy levels. The loss value has oscillatory nature for training and uniform values for the validation. In 10% of edge cropping, accuracy values for both the training and validation of the model are changed over the 100 epochs. There has been a significant improvement in the training accuracy of the model. The loss value has been increased but becomes stable only after 30 epochs. By applying 15% of edge cropping, the accuracy and loss values for both the training and validation of the model do not improve noticeably over the 100 epochs. While applying 20% of edge cropping, we noticed that the accuracy and loss values for both the training and validation of the model have improved over the 100 epochs but the values are in the higher range in comparison with the other cropping ranges. The loss values stay nearly constant over all the epochs.

To choose the best option among these cropping options, accuracy and loss values are listed as shown in Table 4. Table data reflect the readings of the corresponding parameters after 100 epochs. Among the four options, it is observed from both the tabular values and trends on accuracy and loss plots that the 10% cropping has better accuracy and loss values. In other words, considering higher values of accuracies (0.76 and 0.65 for training and validation), lower value of loss (0.4 for training and 0.8 for validation) and curve trends over 100 epochs, 10% cropping achieves the best results.

It has also been observed from this result analysis that increasing edge cropping from 5 to 10% results in increased noise removal. Considering the context of an image, this leads to increased accuracy. On the other hand, increasing edge removal from 15 to 20% has shown a decrement in performance in terms of lower accuracy and higher losses. For the image classification of gully and drainage images, 10% edge

image cropping is selected as it shows the best performance among the alternatives.

4.3.1 Binary classification vs categorical classification

Analysing the gully and drainage images, it has been observed that there are many images that do not fully qualify for either of the classifications ('blocked' or 'not blocked'). These images better fit in the third condition of blockage called "partially blocked". Hence, the binary classification mode with "blocked" and "no blockage" classes does not reflect the real-world monitoring conditions. Therefore, there is a need for a multi-class mode classifier that also allows classifying gullies and drainage images into "partially blocked" classification. Arguably, considering the small number of training dataset, there are more numbers of images in each class in binary mode than that of categorical mode. So, we have evaluated binary classification and categorical classification with the experimental results by comparing the classifier's accuracies performance of both cases. The experiment results, listed in Table 5, showed that the accuracies of categorical class mode are nearly the same as of the binary class mode classifier. These results demonstrate that there is not much difference in the binary and categorical mode of classification from the perspective of classification accuracy. However, from the performance evaluation and the fact that the categorical class mode reflects closer classifications condition of real-world monitoring of such images, categorical classification mode is a better choice.

4.3.2 Image classification with cropping and without cropping

The classifier has been trained and tested for both the cases: images with cropping and images without cropping. We evaluate the performance of our deep learning model for both datasets. From the plots, as shown in Fig. 7a, b, we observed that there are improvements in the training and validation accuracy and loss for the proposed image cropping-based DCNN image classification. The accuracy levels are increased by 21% and 25% for training and validation for image cropping-based DCNN image classification, respectively, as listed in Table 6. There is a significant improvement in the loss value from 0.9 (model without crop-

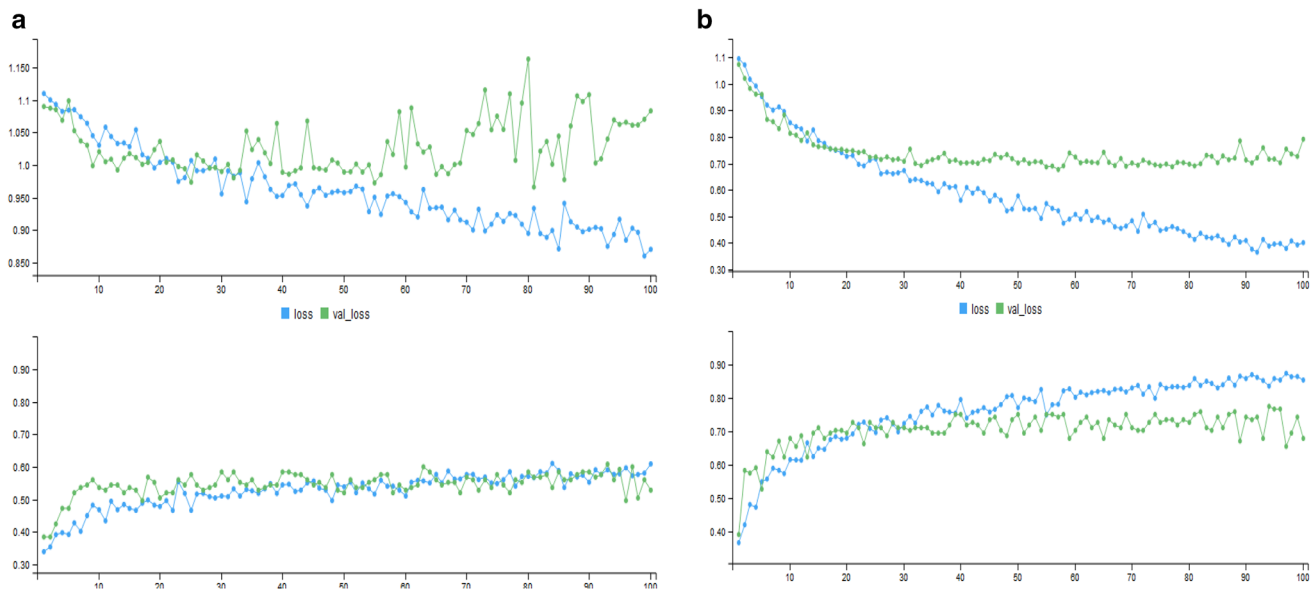


Fig. 7 **a** Accuracy and loss plot on training for model training and validation with images without cropping. **b** Accuracy and loss plot on training for model training and validation with images with cropping

Table 6 Result summaries of training, validation and test of the model

Parameter	Accuracy		Loss	
	Without cropping	With cropping	Without cropping	With cropping
Training	0.55	0.76	0.9	0.4
Validation	0.45	0.7	1.1	0.85
Test	0.6	0.66		

Fig. 8 Correctly classified examples



ping) to 0.4 (model with cropping) during training. Also, there is an improvement in loss value for the validation of the model. These improvements are decisive on classifying images into their corresponding classes. Since lower the loss value, better the classification model, image cropping-based classifier showed better performance.

The model has been tested with randomly sampled images. Figures 8 and 9 show some of the examples of the test results on the classification of images using our model. Images in Fig. 8 are the example of correctly classified images whereas images in Fig. 9 are the examples where images are not correctly classified. From these results, it is observed that the model correctly classified those images which have characteristic features closer to their corresponding classes as it can be observed visually.

4.3.3 Comparison with another classification approach

The major focus of the work reported in this paper is to present a novel use of image classification in identifying gully and drainage blockages for flooding events in a flood monitoring application. In doing so, we have validated our classifier with another baseline classifier considering limitations associated with the small training dataset. Following the implementation of our deep learning-based image classifier for gully and drainage images, the classifier performance is compared with SVM classifier performance on the same image dataset. The experimental results, as listed in Table 7, showed that our proposed model has better accuracy for both training and test results. Our model outperforms the SVM classifier with the increments of 11% accuracy for train-

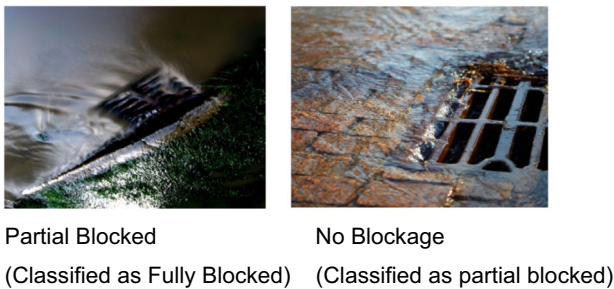


Fig. 9 Incorrectly classified examples

Table 7 Performance comparison between deep learning with cropping image classification and SVM classification

Model/parameter	Accuracy (training)	Accuracy (test)
SVM	0.67	0.62
Deep learning with cropping	0.76	0.66

ing dataset and 4% accuracy increments for the test dataset. These results validated that the deep learning-based image classifier can be used for the purpose of classifying gully images.

5 Conclusions and future work

In our research, as part of a flood monitoring use case, we applied image cropping-based DCNN model for image classification tasks that classify gully and drainage images into either of three blockage classes based on severity: “fully blocked”, “partially blocked” or “no blockage”. At the data pre-processing phase, we applied pixel-wise 10% edge cropping from all four sides of the image to localise the centrally focused region of interest within images. The reduced image size dataset, after applying image cropping, was used for the training and validation of the model. We performed a comparative analysis between categorical and binary classification mode classification performances. The comparison analysis showed that our proposed multi-label image classification model has similar accuracy to that of binary classification. However, considering the multi-category classification in real-time monitoring scenarios as in our use case, the multi-label classification is highly effective to indicate the severity of blockages. Our model has been trained with a small number of publicly available drain blockage images and videos since no publicly available image dataset exists for this use case domain. The experimental results indicate that image classification using image cropping applied at the data pre-processing stage has higher accuracy in comparison to classification without cropping. This classification model also demonstrated an increased accuracy in comparison with an SVM classifier. This proposed image

classification approach for real-time monitoring and identification of flooding events based on gully and drainage blockages has enormous potential as an alternative to either the sensor-based systems or the traditional manual visual inspection, both options being highly expensive and often infeasible.

This work has been performed with a small number of the crowd-sourced image dataset. As part of the future work, a finer analysis of the level of blockage will be explored with a larger dataset diverse in context, environment, and types of objects within images. With such dataset, the accuracy of the classification can be increased which will further enhance the scope of this work to real-time incident monitoring. Also, with the availability of more diverse images, this work can further be analysed with more than three labels of classification. Apart from these aspects, the context and identification of types of objects within the image will additionally help to make a more accurate model with multi-label image classification. The improved classification model will be utilised in an IoT-enabled camera and will act as a real-time monitoring sensor as a part of our future work.

Acknowledgements The research presented in this paper is funded by the European Regional Development Fund Interreg project Smart Cities and Open Data REuse (SCORE).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Chen Y, Lin Z, Zhao X, Wang G, Gu Y (2014) Deep learning-based classification of hyperspectral data. *IEEE J Select Top Appl Earth Obs Remote Sens* 7(6):2094–2107
2. Lee H, Kwon H (2017) Going deeper with contextual CNN for hyperspectral image classification. *IEEE Trans Image Process* 26(10):4843–4855
3. Liu P, Zhang H, Eom KB (2017) Active deep learning for classification of hyperspectral images. *IEEE J Select Top Appl Earth Obs Remote Sens* 10(2):712–724
4. Hu W, Huang Y, Wei L, Zhang F, Li H (2015) Deep convolutional neural networks for hyperspectral image classification. *J Sens* 2015:12 (Art. no. 258619)
5. Chen Y, Jiang H, Li C, Jia X, Ghamisi P (2016) Deep feature extraction and classification of hyperspectral images based on convolutional neural networks. *IEEE Trans Geosci Remote Sens* 54(10):6232–6251

6. Bengio Y, Courville A, Vincent P (2013) Representation learning: a review and new perspectives. *IEEE Trans Pattern Anal Mach Intell* 35(8):1798–1828
7. Hinton GE, Salakhutdinov RR (2006) Reducing the dimensionality of data with neural networks. *Science* 313(5786):504
8. Deng J, Dong W, Socher R, Li L, Kai L, Li F-F (2009) ImageNet: a large-scale hierarchical image database. In: 2009 IEEE conference on computer vision and pattern recognition, pp 248–255
9. Krizhevsky ISA, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. In: Proceedings of the advances in neural information processing systems 25 (NIPS '12), pp 1097–1105
10. Foody GM, Mathur A (2004) A relative evaluation of multi-class image classification by support vector machines. *IEEE Trans Geosci Remote Sens* 42(6):1335–1343
11. Su Y, Jiu-Fen Z, Jiu-Ling Z, Qing-Zhen L (2008) A method for image classification based on kernel PCA. In: 2008 international conference on machine learning and cybernetics, vol 2, pp 718–722
12. Ding C, He X (2004) K-means clustering via principal component analysis. In: Presented at the proceedings of the twenty-first international conference on Machine learning, Banff, Alberta
13. Wang X, Kong Y, Gao Y, Cheng Y (2017) Dimensionality reduction for hyperspectral data based on pairwise constraint discriminative analysis and nonnegative sparse divergence. *IEEE J Sel Top Appl Earth Obs Remote Sens* 10(4):1552–1562
14. Guillamet D, Vitrià J, Schiele B (2003) Introducing a weighted non-negative matrix factorization for image classification. *Pattern Recognit Lett* 24(14):2447–2454
15. Bruce LM, Koger CH, Jiang L (2002) Dimensionality reduction of hyperspectral data using discrete wavelet transform feature extraction. *IEEE Trans Geosci Remote Sens* 40(10):2331–2338
16. Samadzadegan F, Hasani H, Schenk T (2012) Simultaneous feature selection and SVM parameter determination in classification of hyperspectral imagery using ant colony optimization. *Can J Remote Sens* 38(2):139–156
17. Zhao W, Du S (2016) Spectral–spatial feature extraction for hyperspectral image classification: a dimension reduction and deep learning approach. *IEEE Trans Geosci Remote Sens* 54(8):4544–4554
18. Yang C-C et al (2003) Application of decision tree technology for image classification using remote sensing data. *Agric Syst* 76(3):1101–1117
19. Bosch A, Zisserman A, Munoz X (2007) Image classification using random forests and ferns. In: 2007 IEEE 11th international conference on computer vision, pp 1–8
20. Du P, Samat A, Waske B, Liu S, Li Z (2015) Random Forest and Rotation Forest for fully polarized SAR image classification using polarimetric and spatial features. *ISPRS J Photogramm Remote Sens* 105:38–53
21. Huang K, Li S, Kang X, Fang L (2015) Spectral–spatial hyperspectral image classification based on KNN. *Sens Imaging* 17(1):1
22. Amato G, Falchi F (2010) kNN based image classification relying on local feature similarity. In: Presented at the proceedings of the third international conference on SImlarity Search and APplications, Istanbul
23. Tzotsos A, Argialas D (2008) Support vector machine classification for object-based image analysis. In: Blaschke T, Lang S, Hay GJ (eds) Object-based image analysis: spatial concepts for knowledge-driven remote sensing applications. Springer, Berlin, pp 663–677
24. Maulik U, Chakraborty D (2017) Remote sensing image classification: a survey of support-vector-machine-based advanced techniques. *IEEE Geosci Remote Sens Mag* 5(1):33–52
25. Millard K, Richardson M (2015) On the importance of training data sample selection in random forest image classification: a case study in Peatland ecosystem mapping. *Remote Sens* 7(7):8489–8515
26. Foody GM, Mathur A (2006) The use of small training sets containing mixed pixels for accurate hard image classification: training on mixed spectral responses for classification by a SVM. *Remote Sens Environ* 103(2):179–189
27. Blaschke T (2010) Object based image analysis for remote sensing. *ISPRS J Photogramm Remote Sens* 65(1):2–16
28. Yan G, Mas JF, Maathuis BHP, Xiangmin Z, Van Dijk PM (2006) Comparison of pixel-based and object-oriented image classification approaches—a case study in a coal fire area, Wuda, Inner Mongolia, China. *Int J Remote Sens* 27(18):4039–4055
29. Duro DC, Franklin SE, Dubé MG (2012) A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. *Remote Sens Environ* 118:259–272
30. Boutell MR, Luo J, Shen X, Brown CM (2004) Learning multi-label scene classification. *Pattern Recognit* 37(9):1757–1771
31. Cheng G, Han J, Zhou P, Guo L (2014) Multi-class geospatial object detection and geographic image classification based on collection of part detectors. *ISPRS J Photogramm Remote Sens* 98:119–132
32. Blaschke T (2003) Object-based contextual image classification built on image segmentation. In: IEEE workshop on advances in techniques for analysis of remotely sensed data, pp 113–119
33. Li L-J, Su H, Lim Y, Fei-Fei L (2012) Objects as attributes for scene classification. Springer, Berlin, pp 57–69
34. Zhong P, Gong Z, Li S, Schönlieb C (2017) Learning to diversify deep belief networks for hyperspectral image classification. *IEEE Trans Geosci Remote Sens* 55(6):3516–3530
35. Giacinto G, Roli F (2001) Design of effective neural network ensembles for image classification purposes. *Image Vis Comput* 19(9):699–707
36. Chen Y, Zhao X, Jia X (2015) Spectral–spatial classification of hyperspectral data based on deep belief network. *IEEE J Sel Top Appl Earth Obs Remote Sens* 8(6):2381–2392
37. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: 2016 IEEE conference on computer vision and pattern recognition (CVPR), pp 770–778
38. Li T, Zhang J, Zhang Y (2014) Classification of hyperspectral image based on deep belief networks. In: 2014 IEEE international conference on image processing (ICIP), pp 5132–5136
39. Mortensen TM (2014) Blurry and centered or clear and balanced? *Journal Pract* 8(6):704–725
40. Orabona F, Metta G, Sandini G (2005) Object-based visual attention: a model for a behaving robot. In: 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)-workshops, pp 89–89

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.